Evolutionary Journal of the Linnean Society, 2025, **4**, kzaf004
https://doi.org/10.1093/evolinnean/kzaf004
Advance access publication 6 February 2025
Original Article

OXFORD

Original Article

The evolution of herbivory, not terrestrialization, drove morphological change in the mandibles of Palaeozoic tetrapods

Harry O. Berks[1], Pablo S. Milla Carmona[1], Philip C.J. Donoghue[1,*], Emily J. Rayfield[1,*]

[1]Bristol Palaeobiology Group, University of Bristol, Life Sciences Building, Bristol, BS8 1TQ, United Kingdom

[*]Corresponding authors. Bristol Palaeobiology Group, University of Bristol, Life Sciences Building, Bristol, BS8 1TQ, United Kingdom.
E-mail: phil.donoghue@bristol.ac.uk, e.j.rayfield@bristol.ac.uk

ABSTRACT

The radiation of tetrapods during the Devonian and Early Carboniferous was associated with a transition from aquatic to terrestrial environments, with attendant changes in feeding ecology. Despite this, evidence suggests that feeding morphology remained relatively static throughout this transition, until morphological disparity eventually rose later in the Carboniferous and Permian. Using a theoretical morphospace and functional optimality approach, we characterize the functional evolution of tetrapod mandibles, finding an antagonistic relationship between the strength, rotational efficiency, mechanical advantage, and height of jaw morphologies. We further show that the regions of morphospace occupied by the jaws of aquatic and faunivorous terrestrial tetrapods are optimized within this trade-off. As terrestrial herbivores radiated, they explored broader regions of jaw morphospace characterized by deeper, stronger jaw shapes, driving the delayed spike in jaw disparity. We interpret this as a release of functional constraint on jaw morphology by the evolution of herbivory, with new functional demands driving evolutionary innovation. While feeding in aquatic and terrestrial environments is fundamentally different, the criteria for functional optimality in the lower jaw did not change across this transition. Instead, access to terrestrial plant-based diets drove mandibular change.

Keywords: Tetrapoda; theoretical morphospace; functional constraints; optimisation; water-to-land transition

INTRODUCTION

The origin of tetrapods was a formative episode in vertebrate evolutionary history. The shift from aquatic to terrestrial ecology precipitated profound changes in anatomy, including feeding ecology (Long and Gordon 2004, Ahlberg *et al.* 2005, Callier *et al.* 2009, Anderson *et al.* 2011, 2013, Fortuny *et al.* 2011, Heiss *et al.* 2018). The intraoral transport of food is heavily dependent on the fluid medium, leading to differences in behaviour and hyobranchial apparatus between aquatic and terrestrial taxa, and within taxa that feed amphibiously (Stayton 2011, Clack 2012, Cucherousset *et al.* 2012, Heiss *et al.* 2013, 2018, Van Wassenbergh 2013, Witzmann 2013, Michel *et al.* 2015a, 2015b, Van Wassenbergh *et al.* 2017). Terrestrialization would have rendered suction feeding obsolete (Van Wassenbergh *et al.* 2009, Heiss *et al.* 2018), forcing replacement of suction with new mechanisms for food capture, transport and swallowing (Heiss *et al.* 2018) and, in the late Carboniferous, the evolution of terrestrial herbivory (Sues and Reisz 1998). These changes in feeding mode have been linked to morpho-functional changes in the skull and appendicular skeleton in both extant and fossil taxa (Daeschler *et al.* 2006, Van Wassenbergh *et al.* 2006, 2017, Markey and Marshall 2007, Van Wassenbergh 2013, Michel *et al.* 2014, 2015a, 2015b, Neenan *et al.* 2014, Porro *et al.* 2015b, Heiss *et al.* 2018, Lemberg *et al.* 2021). A dichotomy therefore exists between the functional demands of aquatic and terrestrial feeding.

Despite these dramatic changes in environment, it has been shown that the morphological disparity of early tetrapod lower jaws remained fairly static during this transition, peaking later in the Carboniferous (Anderson *et al.* 2011, 2013, Esteve-Altava *et al.* 2013, Neenan *et al.* 2014, Rawson *et al.* 2022). This raises two main questions: why did the variety of jaw shape not change following a dramatic shift in feeding environment, and why did disparity eventually spike a few tens of millions years later? One explanation for the former question may be that the functional demands of biting and suction feeding on jaw morphology may

Received April 29 2024; revised January 30 2025; accepted for publication February 4 2025

be similar, and/or that many finned tetrapodomorphs and early tetrapods engaged in both biting and suction (Porro *et al.* 2015a, 2015b, Lemberg *et al.* 2021). As a result, jaw morphologies of these taxa may be constrained to similar regions of morphospace. Here we hypothesize that aquatic and terrestrial faunivorous (nonherbivorous) jaws were subject to strict functional constraint during this time. As such it would be predicted that the regions of functionally optimal morphospace for extinct aquatic and terrestrial faunivorous taxa overlapped. Following this, it has been noted that the eventual diversification of tetrapod jaw morphologies in the Palaeozoic coincides with the evolution of herbivory, and that herbivory only evolved in crown amniotes and diadectomorphs (Hotton *et al.* 1997, Sues and Reisz 1998, Janis and Keller 2001, Anderson *et al.* 2013). Therefore, we hypothesize that the eventual spike in jaw shape disparity was caused by the evolution of herbivorous jaws, which in turn was delayed by the radiation of diadectomorphs and amniotes.

The phylogenetic limitation of herbivory has inspired many hypotheses linking amniotic synapomorphies and the ability to process plant-based diets (Hotton *et al.* 1997). While this may appear overly simplistic, the scarcity of herbivorous anamniotes suggests the presence of some type of constraint. One such hypothesis proposes that the evolution of herbivory was permitted by the evolution of amniotic jaw muscle conditions, which had more freedom of orientation (Janis and Keller 2001). Anamniotic jaw muscles were limited to anterior orientations from their jaw attachment site, supposedly to maintain sufficient gape with limited skull height (Frazzetta 1968). These muscle orientations have been shown to be effective for kinetic inertial (KI) feeding strategies, where the jaw is accelerated from an open position to a high inertia, to generate enough impulse on closure to kill and process food items (Olson 1961). With the evolution of deeper amniotic skulls, muscles had more freedom to orient posteriorly, allowing for more efficient static pressure (SP) feeding strategies, where the maximum force is applied when the jaw is closed for more effective chewing. It is then argued that SP feeding is a prior requirement for processing plant material and, therefore, herbivorous diets (Janis and Keller 2001). Within this hypothesis, it is possible that the change in functional demands associated with muscle orientation and feeding strategy imposed new constraint on jaw morphology, forcing a change in the occupation of jaw morphospace, and these new jaw shapes in turn allowed the evolution of herbivory. Alternatively, the evolution of crown amniotes and diadectomorphs allowed the evolution of herbivory, whether through new feeding musculature or through other means (Hotton *et al.* 1997, Janis and Keller 2001), which drove selection pressure for new jaw morphologies, resulting in exploration of new morphospace.

To address these questions, we characterized the mandibular morphology of 137 species of stem- and early crown-tetrapods (including outgroup nontetrapod sarcopterygians, ranging in age from Devonian to the end Permian), using elliptical Fourier analysis (EFA) of lateral projections. We used temporal range, phylogenetic, and ecological data to identify two peaks in jaw shape disparity through time, and attribute both to the presence of herbivorous tetrapods. We adopted a theoretical morphospace approach to investigate functional constraints acting on these jaw shapes. Theoretical morphospaces can be used to identify geometrically impossible designs and analyse the functional performance of geometrically viable forms that may or may not be realized in the fossil record (Raup and Michelson 1965, McGhee 1991, Polly *et al.* 2016). Similarly, limitations imposed by extinction and phylogenetic history can be observed in the temporal dynamics of morphological disparity (Guillerme *et al.* 2020a, 2020b). Using a combined theoretical morphology, optimality landscape, and phylomorphospace approach, we used proxies to determine the strength, speed, and hydrodynamic efficiency of the jaw across the terrestrialization event (Deakin *et al.* 2022). We also tested the mechanical advantage of theoretical shapes under an anamniotic, anteriorly oriented muscle model, and an amniotic model of muscle orientation, which is only limited to dorsal direction. Finally, we tested the theoretical shapes under three functional trade-off models: (i) an aquatic anamniotic model, which minimizes jaw height and maximizes strength, rotational efficiency, and mechanical advantage under an anamniotic muscle model; (ii) a terrestrial anamniotic model, which maximizes strength, rotational efficiency, and mechanical advantage under an anamniotic muscle model; and (iii) a terrestrial amniotic model, which maximizes strength, rotational efficiency, and mechanical advantage under an amniotic muscle model. We measured the Pareto rank ratio (PRR) of theoretical shapes in each of these three trade-off models to determine if aquatic or terrestrial jaws were indeed optimized for their hypothesized functional trade-off, thus determining if and how function relates to jaw morphological evolution across the vertebrate water to land transition.

We confirm that aquatic anamniotes are tightly constrained under our aquatic model and that regions of optimal morphospace in the first two models overlap considerably. However, we find that terrestrial anamniotes are not tightly constrained under our terrestrial model, but they are tightly constrained under an aquatic model. Finally, we show that herbivorous taxa are weakly constrained by all three models, though they are the most optimized group under our third functional model. We conclude that terrestrial faunivores experienced evolutionary pressure to minimize jaw height and maximize the mechanical efficiency of forward-facing jaw muscles, which forced evolutionary stasis across tetrapod terrestrialization. This stasis was broken by the evolution of herbivores, which drove the evolution of new jaw morphologies. Functional constraints on the jaw did not restrict the evolution of herbivory; instead, the evolution of herbivory relaxed functional constraint on the jaw.

MATERIALS AND METHODS

Empirical morphometrics

Jaw morphology was measured using 600 coordinates sampled from an outline traversal algorithm of 137 lateral images of tetrapod jaws. These were then input into an EFA (Kuhl and Giardina 1982). A total of 12 harmonics were used, which captured over 99% of shape in over 99% of taxa. A standard principal component analysis (PCA) was used to find the three axes of maximum variation for the construction of two- and three-dimensional theoretical morphospaces (Fig. 1). The theoretical morphospace was built from a 10-by-6-by-6 grid of evenly spaced points, covering the range of empirical jaw shapes and an extra 20% margin, within the first three principal component (PC) axes. In total, 360 theoretical shapes were generated

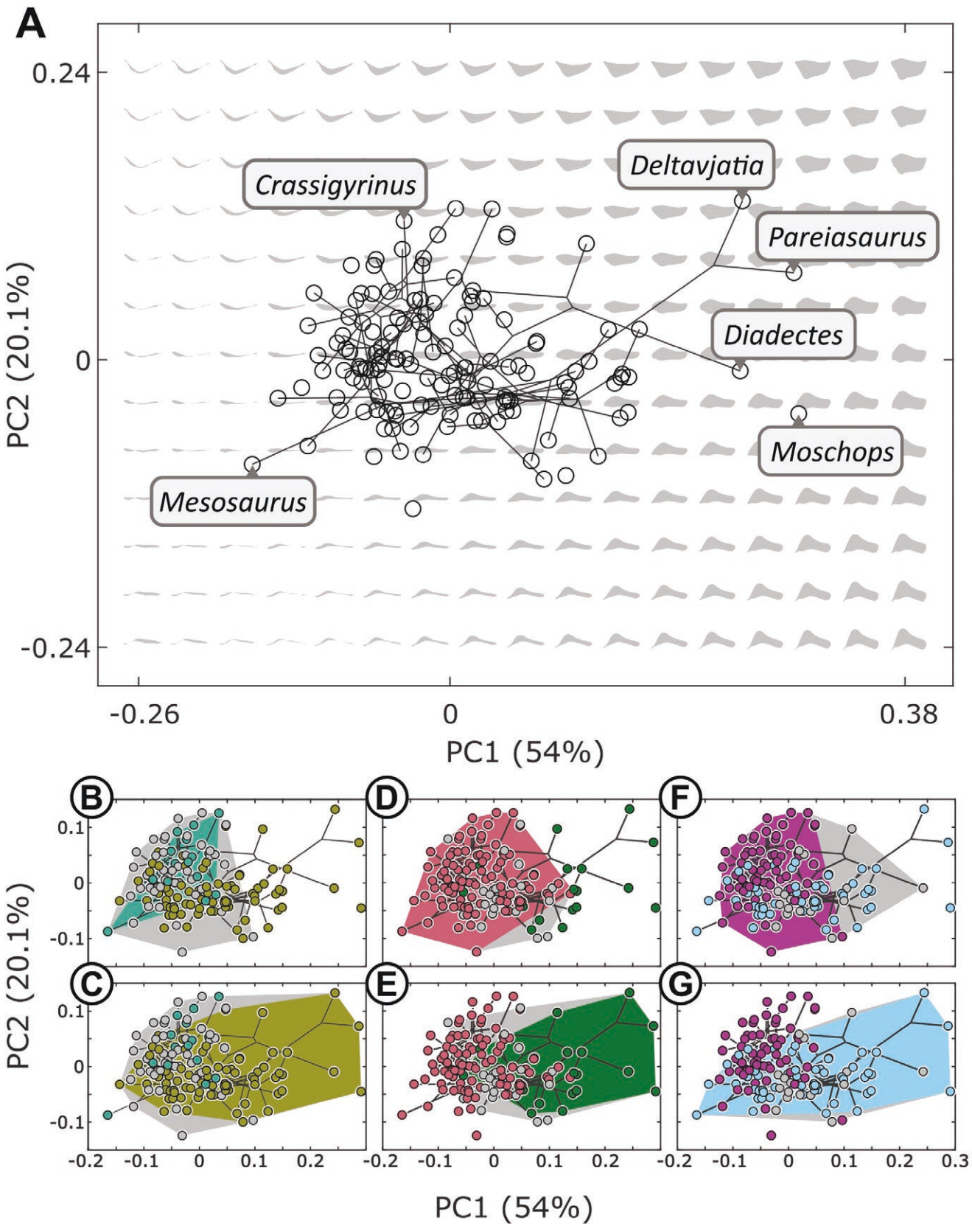

Figure 1. A, early tetrapod jaw empirical phylomorphospace using the temnospondyl hypothesis tree and 2D theoretical morphospace. Theoretical morphospace here is for visualization; a 10-by-6-by-6 3D theoretical grid was used for all functional analyses. Empirical data were split into three binary categories: aquatic (B) and terrestrial (C); nonherbivorous (D) and herbivorous (E); and anamniotic (F) and amniotic (G). Each category has a third uncertain assignment (grey circles). Coloured convex hulls represent the extent of morphospace explored by certain members of a group, and the combination of the coloured and grey convex hulls represents the combined occupation of members of each group with members of the uncertain group.

by transforming from the evenly spaced PC1-2-3 coordinates back to EFA descriptors. In total, 94 of these shapes were self-intersecting loops and therefore identified as impossible shapes, denoting a region of impossible morphospace at low PC1 (McGhee 2015). This left 266 theoretical shapes to test functionally. All morphological analyses were performed with custom MATLAB code, available in the Supporting Information.

Taxonomic data

First and last occurrence dates for taxa were sourced from the literature. The majority of dates were taken from the International Chronostratigraphic Chart (Cohen *et al.* 2013), and a list of sources for each taxon, as well as first and last appearance dates, can be found in the Supporting Information (Table S3).

We assigned each taxon the following categorizations: (i) aquatic or terrestrial; (ii) faunivorous (not herbivorous) or herbivorous; and (iii) anamniotic or amniotic. Each categorical variable was assigned a third 'uncertain' rank, in cases where data were ambiguous or unresolved (e.g. where taxa were identified as semi-aquatic, or there was any controversy in environmental ecology in the literature). The only uncertain amniotes were the recumbirostrans and the diadectomorphs, which have been identified as both amniotes and anamniotes in recent phylogenetic analyses (Pardo *et al.* 2017, Marjanović and Laurin 2019, Reisz *et al.* 2024). Stem amniotes, excluding herbivorous diadectomorphs, were assigned to the anamniotic group. Most ecological groups were based on holistic diagnoses in published literature, which included but was not limited to evidence from limb, body, and tooth morphology, gut contents, and lithology of fossilization. Ecological diagnoses dependent on jaw shape alone were avoided. Some taxa for which we could not find published ecological inferences could be identified as ostensibly faunivorous from tooth morphology. We stress that many of these ecological assignments are conservative; some taxa are coded as uncertain even when they are conventionally interpreted as aquatic, terrestrial, faunivorous, to reflect debate in the literature. Amphibious or semiaquatic taxa have also been included in the 'uncertain' category. A list of all taxonomic assignments as well as the sources used to identify each taxon is available in the Supporting Information.

Temporal dynamics

Taxa were binned into 50, 2.716-Myr bins ranging from the first appearance of the oldest taxon, 387.7 Mya, to the end Permian 251.9 Mya. Bin size can affect results and interpretation in disparity analyses (Guillerme and Cooper 2018), so disparity analyses were performed with 10, 50 and 200 equal bins. These did not change our interpretations, so we only report results corresponding to the 50-bin data. The other signals can be found in Supporting Information Figures S5–S7. Sample sizes of each group were counted within each bin (Fig. 2A, C, E). For each bin, disparity was calculated via the sum of variances, and taxa were bootstrapped 1000 times. Four disparity curves were generated: a total disparity, taking bootstrap samples from all taxa in each bin (Fig. 2B, D, F, grey line) and three group disparities (faunivorous taxa, aquatic taxa, and anamniotes). To take into account the effect of uncertain groups, each group disparity bootstrap sample was generated as follows: (i) uncertain taxa were assigned randomly to one of the other two groups with an equal chance; (ii) the group to be measured was randomly subsampled with replacement, including the extra taxa randomly assigned to it; (iii) disparity was calculated; and (iv) step i–iii were repeated for each bootstrap sample. Deviation between the group disparity curve and the total disparity curve represents the contribution to disparity of the other group; for example, deviation between the red faunivore line and the grey total line represents the amount of total disparity that can be attributed to herbivores. All temporal analyses were performed with custom code written in MATLAB, available in the Supporting Information.

Phylogenetics

We built two informal supertrees in order to account for the phylogenetic uncertainty in early tetrapod phylogenetics, particularly the position of Lissamphibia (Marjanović and Laurin 2019, Kligman *et al.* 2023, San Mauro *et al.* 2023, Marjanović *et al.* 2024). The first uses the phylogeny from Pardo *et al.* (2017) as a backbone resulting in a tree which represents the more commonly supported temnospondyl origin hypothesis and includes 89 taxa from our dataset (Supporting Information Fig. S9). The second tree uses the 'R4' tree from Marjanović and Laurin (2019) as a backbone, representing 110 taxa from our dataset in the lepospondyl hypothesis (Fig. S10). Onto these trees, individual taxa and clades where mapped using trees from multiple sources in the literature (Tsuji and Müller 2009, Swartz 2012, Schoch 2013, Romano and Nicosia 2015, Boos *et al.* 2016, Ford and Benson 2020, Klembara *et al.* 2020b, Reisz *et al.* 2024). These trees were dated using the 'equal' method of the function timePaleoPhy in the R package paleotree (Bapst 2012). For both supertrees, 1000 time trees were generated with random root ages, and mean branch lengths were used for all analyses. These trees were then used to generate ancestral morphologies using the ancestral state reconstruction in the function gm.prcomp in the R package geomorph (Adams *et al.* 2021). This package was also used to test the phylogenetic signal using the function physignal.

Functional analyses

Seven functional performance proxies were calculated for each possible theoretical shape in the theoretical morphospace. These include median von Mises stress (VMS) for strength, rotational efficiency for speed, four types of mechanical advantage for testing the variation in feeding modes and muscle orientation, and, finally, height for hydrodynamic; each of these is explained in greater detail below. From these, function was then inferred at empirical jaw shape PC coordinates to infer the breadth of taxonomic function. All functional metrics relied on aspects of the shape, as well as a set of four algorithmically defined points along its outline, which were used for calculation of functional parameters. We identified positive turning points on each theoretical shape. These are points where the rate of change of the outline slope with movement along the outline reaches a local maximum (i.e. the first derivative of the outline angle is positive and the second derivative of the outline slope is equal to zero) and can be thought of as 'peaks' along the outline (for the formulae and their derivations, see the Supporting Information).

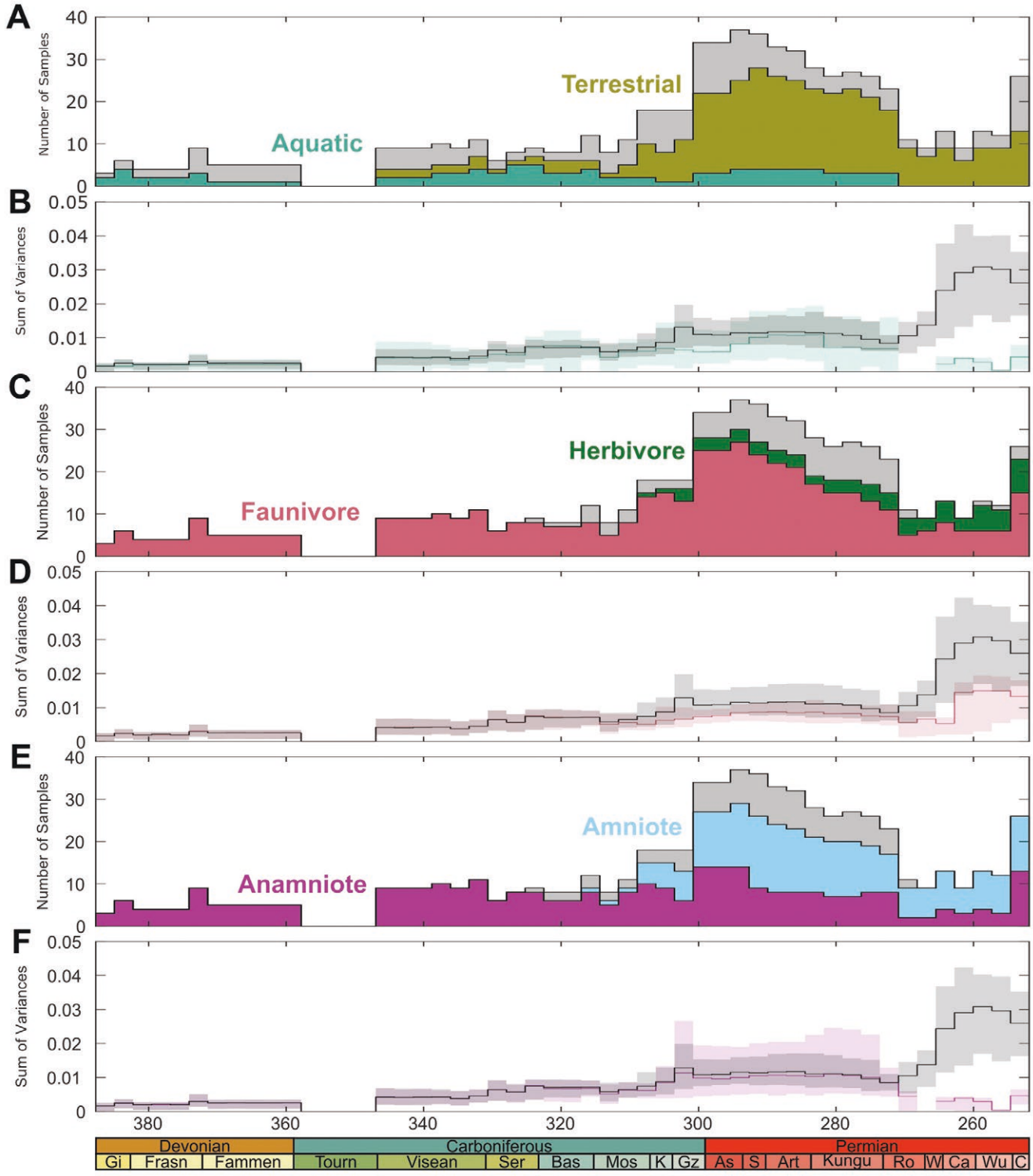

Figure 2. Sample size and disparity through time. Stack plots (A, C and E) represent sample size of each group in each time bin. Grey lines (B, D and F) represent bootstrapped total disparity, and coloured lines represent subsampled disparity [aquatic (blue, B), faunivorous (red, D), and anamniotic taxa (purple, F)]. All subsampled groups have uncertain taxa randomly assigned to them. The difference between the grey and coloured disparity line represents the contribution of the other group to total disparity.

The jaw joint and jaw tip are calculated as the positive turning points on the outline with the lowest and highest x value, respectively. We defined two further points representing the dorsal and ventral extent of possible muscle locations on the jaw. These are located by finding the deepest part of the jaw, or the two points with minimal x coordinate difference that have the greatest difference in their y coordinates. The dorsal and ventral ends of this line are identified as the muscle top and bottom

locations. All points were randomly shifted by 5% of total outline length along the outline, and each functional calculation was performed 1000 times with random input constraints. Results of all performance analyses are shown as performance surfaces in Figure 3. Confidence interval surfaces at the 5% and 95% intervals indicate error, unless otherwise stated, and are depicted in Supporting Information Figure S13.

Strength

Jaw strength is a functional character that has been measured and related to ecology in a wide variety of gnathostomes (Fletcher *et al.* 2010, Gill *et al.* 2014, Neenan *et al.* 2014, Morales-García *et al.* 2019, Coatham *et al.* 2020, Marcé-Nogué *et al.* 2020, Ma *et al.* 2022). The jaw experiences stresses during biting, and we expect these stresses to exert selective pressure on jaw morphology. VMS, a scalar value used to predict when the jaw will yield, of each theoretical jaw shape was measured with a 2D constant strain triangular finite element analysis (FEA), performed with custom MATLAB code. Each shape was standardized to the same area, then automatically meshed to ~2800 elements (varying from 2609 to 3031 elements). The jaw joint and tip positions were used as single node constraints in all degrees of freedom to calculate a global stiffness matrix for the mesh, using an element thickness of 0.01 m, a Young's modulus of 2×10^{10} Pa, and a Poisson ratio of 0.3 (Gill *et al.* 2014, Neenan *et al.* 2014, Morales-García *et al.* 2019). As the global stiffness matrix is very time consuming to compute, the input joint and tip positions were not randomized when calculating VMS in the theoretical jaws. Instead, the top and bottom muscle locations were randomized to apply 1000 pseudorandom force vectors and linearly solve the 1000 von Mises stresses for each element of each model. Muscle orientation was randomized within a range of ±45° from the normal vector at the force node position. Both forces were applied with a strength of 1 N. The quoted VMS value at each point in morphospace is the median of a sample of median VMS values from 1000 finite element model runs. See the Supporting Information for the confidence intervals on these surfaces.

Rotational efficiency

The rotational efficiency (RE) of a jaw shape is defined as in Deakin *et al.* (2022): the speed of the jaw tip given one unit of rotational kinetic energy about its rotational axis (joint). This is derived from the rotational inertia of the structure about the axis, and the length from the axis to the tip (Snively *et al.* 2010, Deakin *et al.* 2022). This was calculated with the same mesh as the VMS calculation, to get the distribution of element masses for calculating the rotational inertia. Input jaw joint and tip constraints were randomized 1000 times by shifting their position randomly within 5% along the output length as stated above. Rotational inertia was calculated with the following formula:

$$I = \sum_{i=1}^{N} m_i r_i^2$$

where N is the number of elements, m_i is the mass of the ith element, and r_i is the distance between the ith element and the rotational axis (i.e. the jaw joint). Following this, the rotational efficiency is calculated with the following formula:

$$RE = L\sqrt{\frac{2}{I}}$$

where L is the distance between the jaw joint and the jaw tip. All functional analyses were performed with custom MATLAB code available in the Supporting Information.

Mechanical advantage

To compare amniote and anamniote muscle conditions, as well as kinetic inertial and static feeding efficiency, four mechanical advantage calculations were performed: (i) kinetic inertial

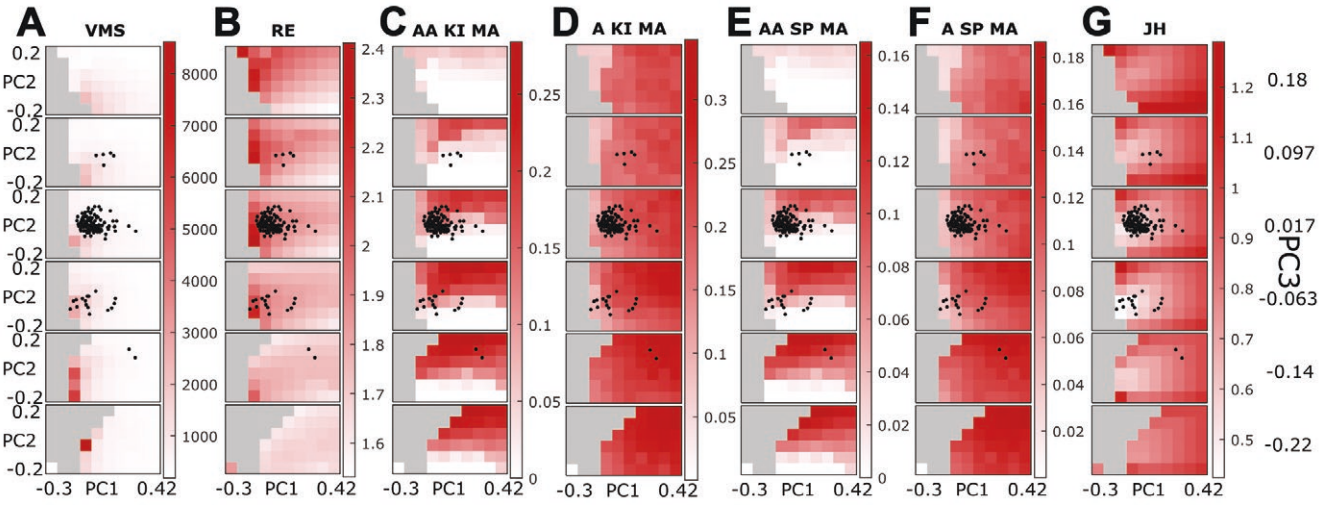

Figure 3. Performance surfaces across the first three principal component axes. Darker colours represent higher values. Grey regions represent impossible morphospace. Black dots represent empirical morphospace. PC1 and PC2 are shown in each surface at a given PC3 value shown on the far right of the figure. Performance variables are: A, von Mises stress (higher values represent higher stress and lower strength); B, rotational efficiency; C, kinetic inertial mechanical advantage of an anamniotic muscle model; D, kinetic inertial mechanical advantage of an amniotic muscle model; E, static pressure mechanical advantage of an anamniotic muscle model; F, static pressure mechanical advantage of an amniotic muscle model; G, jaw height.

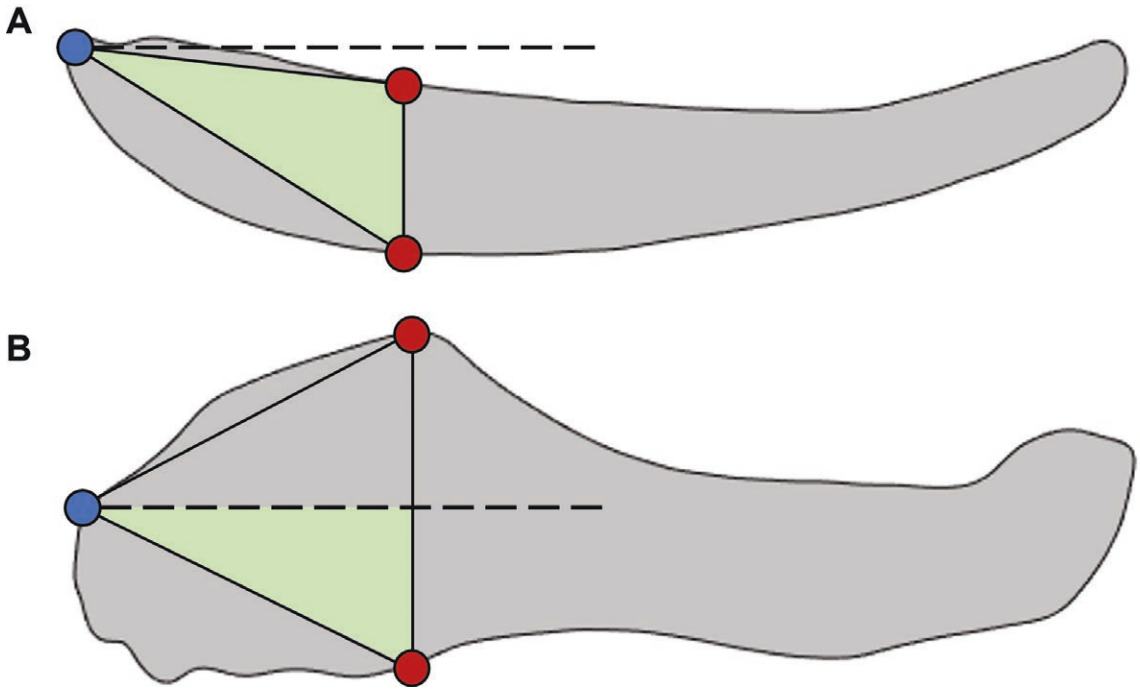

Figure 4. Anamniotic and amniotic muscle models, visualized on an anamniote (*Acanthostega*, A) and an amniote (*Dimetrodon*, B). Blue circle represents jaw joint, and red circles represent muscle top and bottom positions. Dashed line indicates the horizontal line through the jaw joint. Anteriorly oriented anamniotic muscles can effectively be attached anywhere inside the green triangle, below the dashed line. The whole triangle connecting the three circles represents the region effectively utilizable by amniotic muscle conditions with more freedom of orientation.

mechanical advantage of anamniotic muscle orientations; (ii) kinetic inertial mechanical advantage of amniotic muscle orientations; (iii) static pressure mechanical advantage of anamniotic muscle orientations; and (iv) static pressure mechanical advantage of amniotic muscle orientations. The feeding system determines the out-lever, and the muscle condition determines the in-lever.

Kinetic inertial systems rely on accelerating the jaw quickly from an open position. This favours anteriorly oriented muscles but, critically in this case, the output force acts on the centre of mass of the jaw. Static pressure feeding systems act by maximizing the bite force when the jaw is closed. This means that the output force acts at the bite position. Therefore, the kinetic inertial mechanical advantage out-lever is from the jaw joint to the jaw centre of mass, whereas the static pressure mechanical advantage out-lever is from the jaw joint to the jaw tip. This is assuming the animal is biting at the tip, rather than anywhere else along the tooth row. This therefore represents a minimum mechanical advantage, as mechanical advantage is greatest at the posterior end of the tooth row. Amniotes and diadectomorphs differ from anamniotic tetrapods, as they possess deeper pterygoid flanges and much taller skulls which has been a hypothesized driver for the evolution of herbivory, by allowing more efficient static pressure feeding systems (Olson 1961, Janis and Keller 2001). As the skull deepened and the pterygoid flange developed, amniotes had greater freedom to shift the location of their jaw musculature, allowing more muscle variety in position and orientation, facilitating a dual system of adductor mandibulae and pterygoideus muscle arrangement. In anamniotes, the jaw closing muscles are generally oriented anteriorly from their jaw attachment (Frazzetta 1968).

We assumed that the jaw in-lever of a theoretical shape could be placed anywhere along the line between the algorithmically placed muscle top and bottom positions (Fig. 4). Muscle input to the jaw in-lever is maximized when the angle between the in-lever vector and the muscle force is 90°. This means that a vertical muscle force is most efficient when the muscle attachment and the jaw joint lie on a horizontal line. Any muscle oriented anteriorly to the vertical would be most efficient when attached below this horizontal line. Any muscle oriented posteriorly to the vertical would be most efficient when attached above this horizontal line. Figure 4 illustrates how an amniote can make efficient use of all possible in-levers from the jaw joint to any point on the line between the top and bottom muscle points. The total in-lever lengths it can access is characterized by the area of the triangle between the jaw joint, the muscle top point, and the muscle bottom point. In contrast, we assume that an anamniote (excluding diadectomorphs) can only effectively make use of the in-levers below the horizontal line intersecting the jaw joint. This is equivalent to the area of the triangle that sits below the horizontal.

Jaw height

The height of the jaw was measured to identify any functional constraint forced on aquatic taxa by hydrodynamic drag reduction. Here we consider form drag, i.e. drag arising from the shape of an object (as opposed to friction drag due to fluid viscosity). Form drag is related to the size and shape of an object and so

the greater the 'presented area', the higher the drag, while more streamlined objects will have lower drag. This is because the drag forces that aquatic organisms are exposed to when swimming are concentrated on the forward-facing surface (usually, the head) (Nakaya 1995, Van Wassenbergh *et al.* 2015, Gutarra and Rahman 2022). We make two assumptions: (i) that reducing the dimensions of the head relative to flow will lower drag, and (ii) that reduction in dimensions of the head can be approximated by reduction in jaw height. During locomotion, the jaw would usually remain in a closed position in order to maintain a small hydrodynamic profile, so we do not expect the jaw of aquatic organisms to necessarily be streamlined. We do, however, expect selection for streamlined heads to act on both the jaw and skull together, as they combine to make the head. Thus, we predict that nektonic vertebrates will select for shorter jaws (i.e. less tall or with a lower aspect ratio). In contrast, we expect terrestrial organisms to have no preference for jaw height (aspect ratio). Aerodynamic forces on the skull are negligible, and so we expect no unconditional benefit to a shorter (less tall) jaw. We predict that jaw height therefore serves as an important distinction between aquatic and terrestrial jaw morphology.

Jaw height was measured as the maximum y coordinate of the outline minus the minimum y coordinate (i.e. maximum depth). Theoretical jaw outlines were rotated randomly 1000 times within ±45° to measure the error brought by orientation. The minimum values achieved in the 1000 replications are quoted here, with median and maximum values quoted in Supporting Information Figure S10. This metric can be affected by structures such as the coronoid process, which has the potential to disproportionately increase jaw height without the overall head height being affected. However, none of our theoretical jaws shapes—the objects from which height is being measured—have elongate coronoid processes.

Functional optimality

We aimed to test two main hypotheses through functional optimality investigation: (i) the shared jaw morphospace of aquatic and terrestrial faunivorous tetrapods has high functional optimality under both aquatic and terrestrial functional criteria; and (ii) herbivorous jaws extend into a region of morphospace that has high optimality under amniotic functional criteria but not under anamniotic functional criteria. To test the first hypothesis, we generated an aquatic and a terrestrial anamniotic faunivorous Pareto optimality landscape, using the PRR (Deakin *et al.* 2022). Using Pareto optimality, we make no assumptions of weighting or preference for functional criteria, and only assume that aquatic jaws will occupy highly optimal jaw morphospace regions that minimize VMS and jaw height, while maximizing rotational efficiency and mechanical advantage (MA) for kinetic inertial and static pressure feeding strategies under an anamniotic muscle model. We assume that terrestrial jaws will follow the same criteria, but we omit jaw height, which we hypothesize terrestrial anamniotes will have no preference for. We note that, while static pressure feeding systems were not effective in early tetrapod jaw muscle systems (Olson 1961), static pressure force was probably still important in killing and manipulating prey for these organisms. Specifically, under a Pareto framework, high static pressure MA would still convey some functional advantage to kinetic inertial jaws, as any additional force on jaw closure would aid in food processing. For these reasons, our anamniotic conditions consider maximizing anamniotic static pressure MA, as well as kinetic inertial MA. To test the second hypothesis, we generated one more terrestrial amniote jaw optimality landscape (minimize VMS, maximize RE, maximize amniotic kinetic inertial MA, maximize amniotic static pressure MA).

The adaptive landscape metaphor has long been used, with quantitative approaches becoming increasingly common recently, in part because of rapid advancement in computational power (Wright 1932, Arnold 1983, Rayfield 2007, Panagiotopoulou 2009, Svensson and Calsbeek 2012, Polly *et al.* 2016, Ferrón *et al.* 2020). Most studies investigating evolution through this lens have utilized linear weighted sums of functional performance metrics to measure fitness (Arnold 1983, Polly *et al.* 2016, Dickson and Pierce 2019, Stayton 2019a, 2019b, Dickson *et al.* 2021, Jones *et al.* 2021). This proves useful in measuring the relative fitness value of different functional properties. However, this approach requires a priori estimation of the relationship between functional performance and its contribution to fitness and assumes the relationship between function and fitness is continuous. The Pareto optimality approach we use, first proposed by Polly *et al.* (2016), provides a means of testing whether a combination of functional performance metrics is optimized within a sample that is blind to any preference for each functional performance metric, without the need to assign a fitness weighting to each trait. Pareto optimality also assumes that the relationship between optimality and functional performance could be any monotonic function, given an input optimality 'direction' (for each functional performance metric, questioning whether optimality increases or decreases as performance increases). This is still uncertain but is a less strict axiom than continuity.

Pareto optimality, as employed here, is a rank-based system of identifying optimal solutions within any sample of performance combinations for a set of objectives, given the assumption that optimality increases monotonically with each performance metric (Goldberg 1989, Horn *et al.* 1994, Fonseca and Fleming 1998, Alberto *et al.* 2003). The optimal solutions (called the Pareto optimal subset or POS) are defined as the subset of the set of solutions for which no member of the total set possesses a higher or equal performance in all objectives when compared with the subset. Defined mathematically: a solution to any multi-objective optimization problem can be described by its functional performance for each objective. Thus, for a set of N_f functional objectives, each solution can be described by a vector, s_i:

$$s_i = \left(f_{i1}, f_{i2}, \ldots, f_{iN_f}\right)$$

where f_{ij} denotes the functional performance of solution i for functional objective j. Let S_0 be a finite set of solutions:

$$S_0 = \{s_1, s_2, \ldots, s_{N_s}\}$$

where N_s is the number of solutions within S_0. Let P_0 be the subset of all Pareto optimal solutions within S_0. Assuming that optimality increases monotonically with an increase in the performance for any objective, a solution s_i belongs to P_0 if and only if:

$$f_{ij} \geq f_{kj}, \quad \forall j \in \{1, 2, \ldots, N_f\}, \quad \forall k \in \{1, 2, \ldots, N_s\}$$

While all solutions in P_0 can be considered equally optimal, the remaining solutions in S_0 that are not Pareto optimal ($S_0 \quad P_0$) may not all be equally suboptimal, as some are still dominant over others. To further rank the remaining solutions, Goldberg developed an iterative Pareto rank algorithm, where P_0 is removed from S_0 and a new POS is found (Goldberg 1989, Alberto *et al.* 2003). To find the Goldberg rank of a solution, first we define S_n iteratively:

$$S_n = S_{n-1} - P_{n-1}, \quad n > 0$$

It is trivial to prove that if S_n is not empty, then P_n is not empty. Thus, the number of elements in S_n decreases as i increases, until S_N, where $S_N = P_N$. Furthermore, it can be shown that each solution belongs to a unique POS. Let R_O be the Goldberg rank of all solutions within P_{R_O}.

In cases with high heterogeneity in performance space occupation (such as Deakin *et al.* 2022), the Goldberg rank is biased to favour solutions in sparsely occupied regions of performance space. To mitigate this, we use the Pareto rank ratio (*PRR*) from Deakin *et al.* (2022). This is calculated with a second set of Goldberg ranks of the same set of solutions, but with reversed optimality—rather than optimality increasing with increasing functional performance, optimality increases with decreasing functional performance. This provides a suboptimal Goldberg rank, R_S. Let P'_n denote the Pareto suboptimal set of S'_n. A solution s_i within S'_n belongs to P'_n if and only if:

$$f_{ij} \leq f_{kj}, \quad \forall j \in \{1, 2, \ldots, N_f\}, \quad \forall k \in \{1, 2, \ldots, N_{S'_n}\}$$

The ranking process iterates again, but with reversed optimality:

$$S'_n = \begin{cases} S_0, & n = 0 \\ S'_{n-1} - P'_{n-1}, & n > 0 \end{cases}$$

The suboptimal Goldberg rank, R_S, is assigned to all solutions within P'_{R_S}. Finally, the *PRR* is defined as:

$$PRR = \begin{cases} 1, & R_O = 0 \\ \frac{R_S}{R_O + R_S}, & R_O > 0 \end{cases}$$

Thus, the PRR is one when the solution is Pareto optimal, zero when the solution is not Pareto optimal and belongs to the suboptimal front, and some value between zero and one otherwise. Note, this is quoted differently than in Deakin *et al.* (2022), where the nominator of the fraction is decreased by one and the denominator is decreased by two. This is because the Goldberg rank algorithm in Deakin *et al.* (2022) begins at rank one, rather than rank zero.

Functional interpolation

To identify patterns in function and optimality within empirical jaw shapes, we interpolated their functional performance and PRR from their position in jaw morphospace. Specifically, we used a 3D linear interpolation of the theoretical performance and optimality data at each empirical jaw PC1, 2 and 3 coordinate. The distributions of each set of functional metrics and PRRs is shown in Figure 6 below.

RESULTS

Morphological disparity

Empirical morphology shows a high concentration of variance in the first three axes of EFA jaw morphospace (PC1 = 54% of the variance, PC2 = 20%, PC3 = 7.6%). There is considerable overlap in morphospace occupation between the jaws of aquatic and terrestrial taxa, and between the jaws of anamniotes and amniotes (Fig. 1). The jaws of terrestrial taxa encompass a more expansive area of morphospace than do the jaws of aquatic taxa. However, morphospaces extending beyond the aquatic–terrestrial overlap and anamniote–amniote overlap are mostly occupied by herbivores. The distinction between jaws of herbivorous and nonherbivorous taxa is more defined than between other groupings. Herbivorous jaw shapes occupy the region of morphospace characterized by deeper and more robust jaws.

We recover a signal of delayed morphological disparity following terrestrialization across all tetrapods (Fig. 2, grey lines). Specifically, we recover relatively static jaw disparity through time, spanning the appearance of the first definitively terrestrial taxa in the Visean stage, until a slight peak in jaw disparity ~303 Mya in the Gzhelian. Jaw disparity remains static for another ~40 Myr, where it spikes significantly once again at ~265 Mya (Wordian/Capitanian). From this dataset, we informally identify three episodes in the evolution of tetrapod jaws: (i) constrained evolution from the mid-Devonian until the Gzhelian, during which terrestrial taxa begin to evolve; (ii) a spike in disparity in the Gzhelian followed by further stasis until the Capitanian; and (iii) a final spike in the Capitanian, followed by continued static disparity until the end-Permian mass extinction.

The contribution of terrestrial jaws to total disparity in the first episode (the difference between the total grey disparity line and the aquatic blue disparity line in Fig. 2B) is minimal, even after their first definitive appearance in the Visean. The start of the second episode approximates the origin of herbivores and the origin of amniotes. The spike in disparity is contributed entirely by anamniotic terrestrial herbivores, as the disparity in aquatic jaws and faunivorous jaws remains static across this transition, but the disparity in anamniotic jaws follows the total disparity in a spike. Finally, the third episode is characterized by a reduction in sample size and a large spike in disparity. This can be directly attributed to amniotic herbivores, as aquatic, faunivorous, and anamniotic jaw disparity does not spike. Anamniotic jaw disparity falls substantially during this period; the absence from our dataset of aquatic taxa for this time interval is probably a sampling artefact.

Phylogenetic signal

We recover a weak but statistically significant phylogenetic signal in our jaw EFA data regardless of which of our two phylogenies is used (Temnospondyl origin hypothesis: $K_{mult} = 0.5189$, $P = .007$, Lepospondyl origin hypothesis: $K_{mult} = 0.556$, $P = .006$), suggesting that phylogenetic constraint on the evolution of early tetrapod mandibular morphology was present yet did not overwhelm the distribution of taxa in morphospace. Indeed, the phylomorphospace shows high convergence among jaw forms within the dataset, as the branches cross repeatedly (Fig. 1). The amniote clade explores wider regions of morphospace.

Extension of morphospace by herbivores is limited to amniotes and diadectomorphs, as these are the only groups thought to have evolved herbivory (Sues and Reisz 1998). Little difference can be seen in the phylomorphospace when changing between our two phylogenies, suggesting that our results are robust regardless of phylogenetic uncertainty (Supporting Information Fig. S14).

Theoretical morphology, function, and optimality

Theoretical shapes show an increase in jaw depth with increased PC1, and a change from convex to concave jaw curvature with increasing PC2. Increasing PC3 is more obscure, but generally represents increasing posterior jaw depth, and a coincident lowering of the jaw joint. Shapes in low PC1 coordinates across all PC2 and PC3 coordinates were identified as impossible space, with some variation in boundary through PC2 and PC3. Possible theoretical shapes inform a pattern of increasing VMS with decreasing PC1 and PC2, and decreasing PC3. VMS spikes at the boundary between possible and impossible jaw morphospace, where jaw shapes are extremely thin. RE shows a similar pattern of decrease with increasing PC1, but it shows a saddle pattern generated by a peak in the mid-PC2 axis. RE increases with increasing PC3. The distinction between static pressure and kinetic inertial mechanical advantage is minimal. However, the effect of the different amniotic and anamniotic muscle conditions on the mechanical advantage is dramatic. Anamniotic mechanical advantage seems relatively consistent along the PC1 axis but increases dramatically with increasing PC2. It decreases with increasing PC3. Amniotic mechanical advantage shows a stronger positive correlation with PC1, while maintaining the decrease with increasing PC3. Finally, jaw height troughs in central PC1, PC2, and PC3, and increases with distance from the mean in all directions.

We generated three optimality landscapes, each corresponding to a different set of hypothesized optimal functional criteria (see Methods) (Fig. 5). The first model of functional optimality we generated was an aquatic anamniote model. The optimality landscape for this model shows broad regions of optimality in the centre of jaw morphospace, with suboptimal regions limited to extremes of PC1, PC2 and PC3. The second model of terrestrial anamniotic functional optimality is similar, albeit with tighter regions of optimality. The final model of functional optimality, corresponding to terrestrial amniotes, shows further limitations on functionally optimal regions, but with greater optimality with higher PC1 and lower PC3 values, as well as higher PC2 and higher PC3 values.

Interpolated function and optimality

The interpolated functional performance of empirical jaws reveals even distributions between most taxonomic groups and most performance criteria (Fig. 6). The main differences between groups are between herbivorous and nonherbivorous jaws. Herbivorous jaws are stronger, taller, and less rotationally efficient and have higher mechanical advantage under an amniotic muscle model than nonherbivorous jaws. The interpolated optimality of empirical shapes for each of the three functional models varies substantially between groups (Fig. 6H–J). We accept our hypothesis that aquatic jaws are constrained to jaw morphospace that is optimized within our aquatic functional trade-off model. However, terrestrial anamniotes are not optimized within our terrestrial anamniotic functional model. Instead, they are most optimized within our aquatic model. Unexpectedly, we find that nonherbivorous jaws are highly optimized within our aquatic functional trade-off model, with more than 75% of herbivore jaws scoring PRR values >0.5. Herbivores sit in a region of the morphospace that is not tightly optimized under our aquatic and terrestrial anamniotic functional trade-off models, as we expected, but this region has similar PRR values under our amniotic functional trade-off model. They are, however, still quite optimal under all three models with more than half the PRR values >0.5, and they are the most optimal group under the terrestrial amniotic functional model.

DISCUSSION

Theoretical morphospace and optimality

Empirical morphospaces have been utilized extensively to explore the morphological and functional diversity of the jaw (Anderson *et al.* 2011, 2013, Neenan *et al.* 2014, Palci *et al.* 2016, Hill *et al.* 2018, Schaeffer *et al.* 2019). These patterns, while informative, are restricted to morphofunctional relationships seen in biological systems. Furthermore, the variety of shape within biological systems is controlled by many nonfunctional characters (Seilacher 1970). The heterogeneous occupation of empirical morphospaces can be explained through three major constraints: phylogenetic constraints, due to which the limits of morphospace are a function of evolutionary history; morphogenetic constraints, in which certain regions of morphospace are inaccessible to a biological construction process; and functional constraints, where functionally inferior forms are selected against, leaving functionally suboptimal regions of morphospace empty. In contrast, the application of a theoretical morphospace approach allows us to infer shapes that span the region of occupied and unoccupied morphology. Functional tests of these theoretical shapes permit the identification of regions of morphospace that are inaccessible to an evolving sample of empirical shapes, and can provide plausible causal explanations.

Here we have measured the relationship between morphology and function across theoretical jaw shapes that either mirror fossil jaws or describe forms that are unrealized in the fossil record (Raup 1967, McGhee 2006, Polly *et al.* 2016, Deakin *et al.* 2022). We find that the relationships between jaw speed, strength, mechanical advantage, and height are not only constrained by adaptation but also by a fundamental trade-off tied to properties of shape (Fig. 3). Understanding the functional limitations on biological form requires an understanding of the morphological limitations on function—such as antagonistic relationships between functions moderated by shape—which the combination of theoretical morphology and functional optimality approach can achieve. While we focus on functional trade-offs controlling jaw morphospace, we note that the lower PC1 boundary of the occupied jaw EFA morphospace is largely controlled by geometrically inaccessible morphospace, precluding morphological expansion in this direction (McGhee 2006) (Fig. 3). Phylogenetic signal within shape variation is generally weak but significant, suggesting that phylogenetic constraints had some limiting

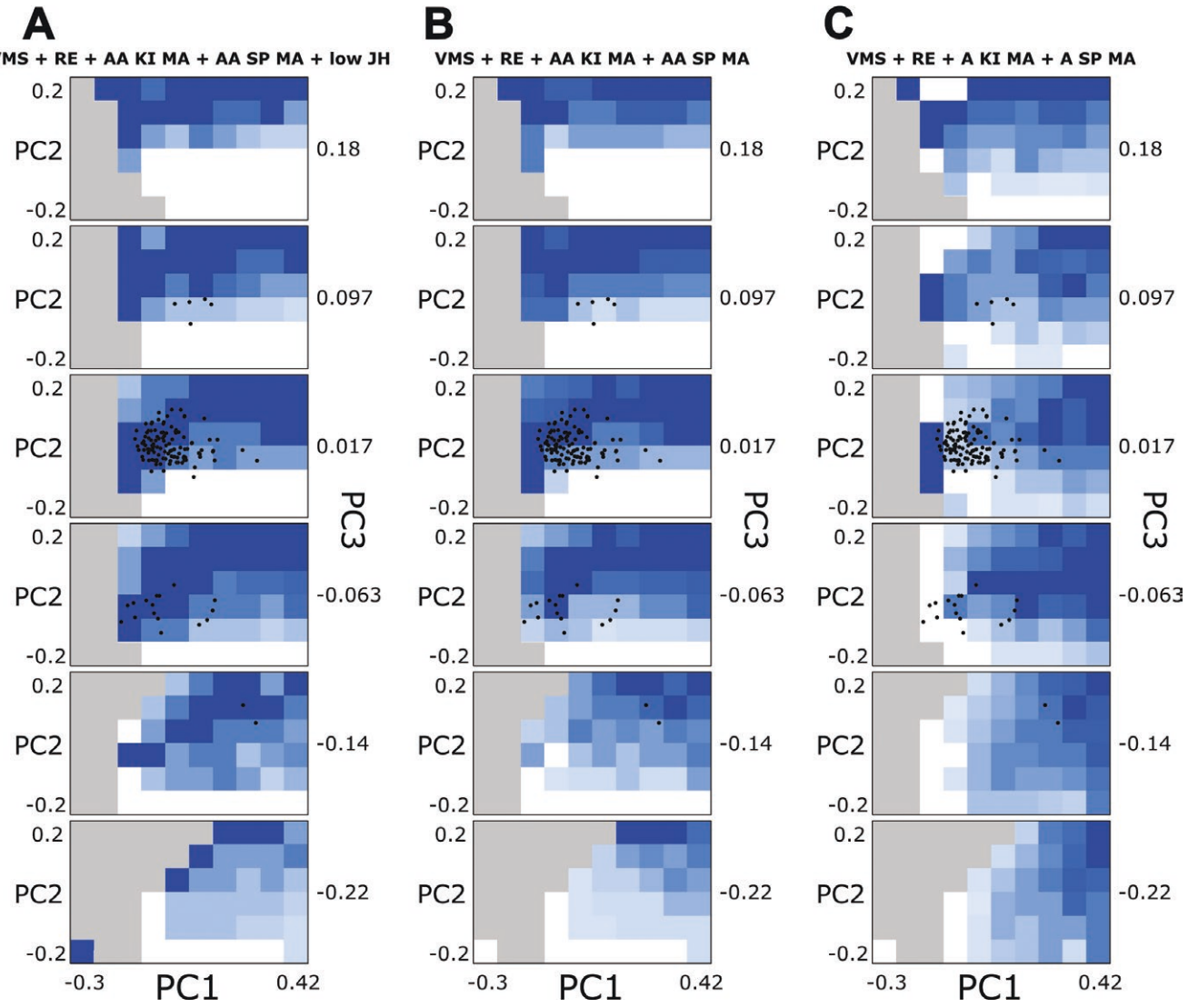

Figure 5. Pareto rank ratio optimality surfaces under the three trade-off models: A, aquatic anamniote model, or the trade-off between minimizing VMS, maximizing RE, maximizing kinetic inertial and static pressure mechanical advantage under an anamniotic muscle model, and minimizing jaw height; B, terrestrial anamniote model, or the trade-off between minimizing VMS, maximizing RE, and maximizing kinetic inertial and static pressure mechanical advantage under an anamniotic muscle model; C, terrestrial amniote model, or the trade-off between minimizing VMS, maximizing RE, and maximizing kinetic inertial and static pressure mechanical advantage under an amniotic muscle model. PC1 and PC2 are shown in each surface at a given PC3 value shown to the right of each trade-off scenario.

effect on morphospace expansion. The functional constraints we identify here work in tandem with geometric and biological constraints to limit early tetrapod jaw morphospace.

Delayed onset of morphological disparity

We identify two distinct peaks in tetrapod jaw disparity in the Gzhelian and the Wordian, which split the temporal morphological dynamics of tetrapod jaws into three phases (Fig. 2). The first episode of relatively static jaw disparity spans the evolution of terrestrialization, corroborating existing conclusions in the literature that the evolution of jaw variety lagged behind other aspects of skeletal anatomy during the initial phase of tetrapod terrestrialization (Anderson *et al.* 2011, 2013, Neenan *et al.* 2014, Rawson *et al.* 2022). To explain this stasis, we hypothesized that functional constraint was strong on these early jaws, and that despite differences in fluid medium and the functional requirements of feeding on land and in water, the optimal regions of jaw morphospace for aquatic and terrestrial anamniotes overlap. While we identified tight functional constraint in aquatic jaws under their hypothesized trade-off, we did not recover the same signal for terrestrial jaws. Instead, terrestrial jaws are also tightly constrained under the aquatic model. In fact, all faunivorous taxa were optimized under the anamniotic aquatic functional model. We conclude that functional constraint was strong during this period of stasis, and that the jaws of Palaeozoic nonherbivorous taxa were under significant evolutionary pressure to reduce their jaw height and maximize the efficiency of anteriorly oriented jaw musculature.

The second episode of static tetrapod jaw disparity occurred in the Gzhelian, defined by a distinct jump in disparity. While it has been concluded from qualitative evidence that this increase in disparity was driven by the evolution of herbivorous

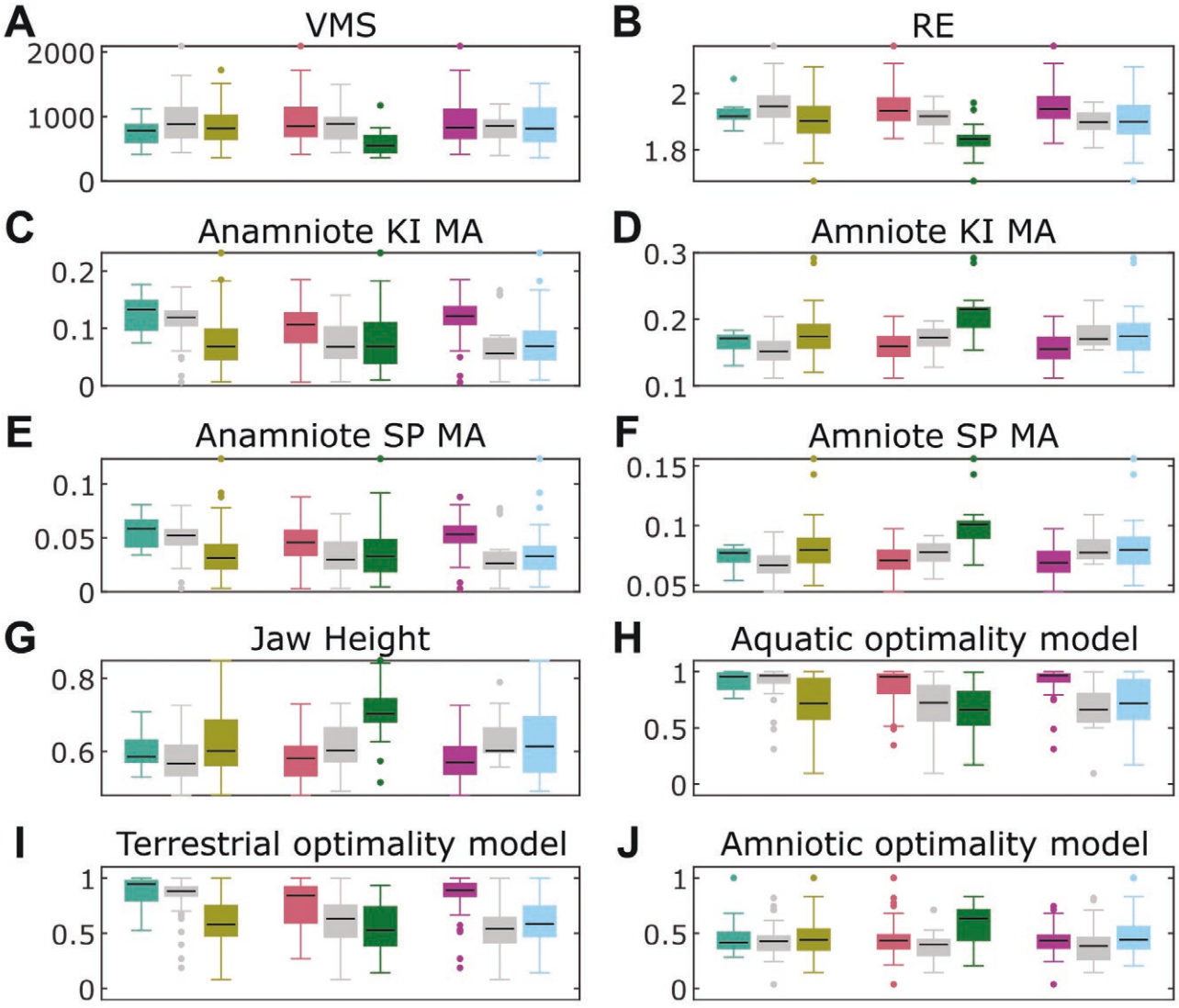

Figure 6. Distributions of interpolated functions and optimalities of empirical jaw shapes. Taxa are split by environment [aquatic (teal), uncertain (grey), and terrestrial (olive green)], diet [nonherbivorous (red), uncertain (grey), and herbivorous (green)], and clade [anamniotic (purple), uncertain (grey), and amniotic (light blue)].

taxa (Anderson *et al.* 2013), we provide the first quantitative proof of this hypothesis, as this spike in disparity is not present when herbivorous taxa are removed (Fig. 2D). Herbivorous jaw morphospace is more optimal than expected under our amniotic functional trade-off model, but it is also relatively optimal under the anamniotic terrestrial and aquatic models (Figs 3, 6). We therefore reject the hypothesis that release of jaw shape functional constraint allowed herbivory to evolve. Instead, the evolution of herbivory was permitted by other means, which in turn drove the change in jaw shape.

Finally, the third episode of Palaeozoic tetrapod jaw disparity began in the Wordian. We note that this follows a decrease in sampling within our dataset at the start of the Roadian, which can be best seen in our high-resolution disparity curve (Supporting Information Figs S5–S7). We therefore attribute this pattern to morphospace unpacking, with loss of data clearing intermediate regions of morphospace occupation. We hypothesize that this loss in sample size and gain in disparity is associated with the prolonged Kungurian/Roadian extinction event (Didier and Laurin 2021, 2024). As morphological variety is maintained with fewer taxa, the sum of variances increases. Notably, this peak can again be attributed to herbivorous jaw shapes.

Herbivory drives jaw morphology

The extremes of the extended terrestrial morphospace region are occupied by four taxa: the diadectomorph *Diadectes*; the pareiasaurs *Deltavjatia* and *Pareiasaurus*; and the therapsid *Moschops* (Fig. 1A). These taxa all show convergent anatomical features. Specifically, they share cranial pachyostosis, and large, stocky, upright appendicular skeletons (Case 1911, Barghusen 1975, Berman *et al.* 1992, 1998, Lee 1997, Tsuji 2013, Benoit *et al.* 2017, Canoville and Chinsamy 2017, Boitsova *et al.* 2019, Klembara *et al.* 2020a, 2020b, Van den Brandt *et al.* 2023). This evidence, combined with tooth morphology and gut contents, has supported interpretations of these organisms as herbivorous (Barghusen 1975, Karlsruhe and Sues 1993, Sues and

Reisz 1998, Eberth *et al.* 2000, Canoville *et al.* 2014). Combined with our morphological data, these data show that herbivorous jaws are fundamentally different from nonherbivorous jaws. We identify that this extension of jaw morphospace is linked to morphospace regions with high mechanical advantage, low rotational efficiency, and low stress, suggesting that the evolution of these shapes was driven by functional demands of herbivory (Fig. 3). Alternatively, this dramatic shift in jaw shape may be due to tight integration between the skull and jaw within these taxa. This would result in jaw morphological convergence driven by the convergently developed cranial adaptations within each lineage (Berman *et al.* 1992, Tsuji 2013). Patterns of increased constraint between skeletal modules has been demonstrated in the skull throughout this period, but the covariation between the skull and jaw remains uninvestigated (Morris and Abzhanov 2021, Rawson *et al.* 2022).

Morphodynamic constraints

While the jaws of aquatic taxa and terrestrial faunivorous taxa are constrained to optimal performance for our aquatic functional model, they do not explore all Pareto optimal regions of morphospace. There may be some further constraint on their morphology not captured by our functional model, related to phylogenetic or morphogenetic factors. While the phylogenetic signal in these morphological data is weak, wider areas of morphospace are occupied by amniotes and some diadectomorphs (Fig. 1). However, amniotes in this dataset are almost entirely terrestrial and so untangling the phylogenetic and functional biases on jaw form is difficult. The aquatic amniote *Mesosaurus* lies beyond the range of nonamniotic aquatic taxa, suggesting that limits on jaw shape variation in anamniotes may be phylogenetic (Fig. 1). However, *Mesosaurus* also does not lie in the extended region of terrestrial morphospace. *Mesosaurus* has been suggested to be a filter-feeder (Villamil *et al.* 2016), which could explain its separation from the rest of the dataset. Still, areas of Pareto optimal morphospace remain unoccupied by all taxa within this dataset. This could be due to thresholds acting on the individual performance metrics—for example, there may be a minimum strength the jaw must exceed in order to perform. Theoretical morphologies below this threshold can still be considered optimal if they have high rotational efficiency or a low jaw height (Deakin *et al.* 2022).

Morphogenetic constraints may also play a role in the constraint of nonamniotic tetrapod jaws. Metamorphosis has been characterized in nonamniotic tetrapod fossils from this period, with the majority of taxa exhibiting aquatic larval forms (Schoch 2009). Aquatic larval stages can act to constrain and promote diversification in adults (Powder *et al.* 2015). However, studies on salamander morphology have shown that metamorphosis has promoted morphological diversification in bones associated with feeding (Fabre *et al.* 2020). Instead, the limitation may arise from complex life cycle shifts during metamorphosis, between aquatic and terrestrial environments (Bonett and Blair 2017). Furthermore, the need to return to water for reproduction may strongly enforce these constraints, which provides an explanation for the similar constraint acting on both aquatic and semiaquatic jaws. The evolution of the amniotic egg would thus relieve this constraint by allowing fully terrestrial juvenile forms, providing an explanation as to why some amniotes cross this morphospace boundary and explore greater PC1 regions.

Alternatively, this constraint may have been lifted by the evolution of a fully terrestrial feeding system, with a mobile neck and the remodelling of the ancestral hyobranchial system (Witzmann 2013). Some aquatic taxa can feed on terrestrial food sources by either carrying them back to the water, or pumping water through the mouth to aid in intraoral transport of food (Stayton 2011, Cucherousset *et al.* 2012, Heiss *et al.* 2013, 2018, Van Wassenbergh 2013, Michel *et al.* 2015b, Van Wassenbergh *et al.* 2017). With the evolution of mobile necks and a prehensile tongue, the jaw may be alleviated from the evolutionary pressure to pump water efficiently. Future studies linking the morphology of the hyobranchial apparatus with jaw optimality can test this hypothesis.

CONCLUSION

Our aim here was four-fold: (i) to recover the signal of delayed morphological innovation in terrestrial tetrapod jaws with morphometric analysis; (ii) to identify functional constraints acting on aquatic and early terrestrial tetrapod jaws that may explain this stasis; (iii) to quantitatively test whether this delayed morphological spike was driven by the evolution of herbivory; and (iv) to test the hypothesis that the delayed evolution of early tetrapod herbivores could be attributed to jaw functional constraint imposed by anamniotic anteriorly oriented jaw muscle arrangements (Janis and Keller 2001). We achieved our first aim, showing two distinct peaks in tetrapod jaw disparity through the Palaeozoic. We further show that aquatic jaws are tightly constrained to minimize stress, maximize rotational efficiency, maximize mechanical advantage, and minimize jaw height. Unexpectedly, terrestrial faunivores also adhere to these constraints. However, large optimal areas under this trade-off remained unexplored, suggesting that the limiting factor on jaw disparity was either nonfunctional, or the relative weighting preference for functional traits under an additive fitness model (Arnold 2003, Polly *et al.* 2016). We demonstrate that the eventual spike in morphological disparity can be attributed to the evolution of herbivory. However, we can reject the hypothesis that released jaw constraint allowed herbivory to evolve. Instead, the evolution of herbivory was permitted by other means, whether the evolution of the muscular system itself, or larger guts for the digestion of plant material (Hotton *et al.* 1997). Herbivory then drove rapid diversification of jaws into new regions of morphospace.

SUPPLEMENTARY DATA

Supplementary data are available at *Evolutionary Journal of the Linnean Society* online.

ACKNOWLEDGEMENTS

We express our gratitude to William Deakin who completed an earlier version of this study in partial fulfilment of his MSci Palaeontology and Evolution at the University of Bristol; Will declined authorship of this final version of the study. We thank Tom Smith for helping with the ancestral state reconstruction coding.

CREDIT STATEMENT

Harry Berks (Data curation [Equal], Formal analysis [Equal], Investigation [Equal], Validation [Equal], Writing—review & editing [Equal]), Pablo Milla Carmona (Formal analysis [Equal], Validation [Equal], Writing—review & editing [Equal]), Philip Donoghue (Conceptualization [Equal], Funding acquisition [Equal], Methodology [Equal], Project administration [Equal], Supervision [Equal], Writing—review & editing [Equal]), and Emily Rayfield (Conceptualization [Equal], Funding acquisition [Equal], Project administration [Equal], Supervision [Equal], Writing—review & editing [Equal]).

CONFLICT OF INTEREST

All authors declare that they have no conflicts of interest.

FUNDING

E.J.R. and P.C.J.D. acknowledge the support of the John Templeton Foundation [Grant 62574; the opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation]. P.C.J.D. is also funded by a Leverhulme Trust Research Fellowship (RF-2022-167) and E.J.R. was funded by Standard Grant NE/P013090/1 from the Natural Environment Research Council.

DATA AVAILABILITY

Our jaw shape outline data are archived on Figshare with 10.6084/m9.figshare.27993197, and the code for analysis is available at: https://github.com/Bristol-Palaeobiology/theofun.

REFERENCES

Adams DC, Collyer ML, Kaliontzopoulou A, et al. *Geomorph: Software for Geometric Morphometric Analyses*. R package version 4.0. 2021.

Ahlberg PE, Clack JA, Blom H. The axial skeleton of the Devonian tetrapod *Ichthyostega*. *Nature* 2005;**437**:137–40. https://doi.org/10.1038/nature03893

Alberto I, Azcarate C, Mallor F, *et al*. Multiobjective evolutionary algorithms. Pareto rankings. *Monografias del Semin. Matem. Garcia de Galdeano* 2003;**27**:27–35.

Anderson PSL, Friedman M, Brazeau MD, *et al*. Initial radiation of jaws demonstrated stability despite faunal and environmental change. *Nature* 2011;**476**:206–9. https://doi.org/10.1038/nature10207

Anderson PSL, Friedman M, Ruta M. Late to the table: diversification of tetrapod mandibular biomechanics lagged behind the evolution of terrestriality. *Integrative and Comparitive Biology* 2013;**53**:197–208. https://doi.org/10.1093/icb/ict006

Arnold SJ. Morphology, performance and fitness. *American Zoologist* 1983;**23**:347–61. https://doi.org/10.1093/icb/23.2.347

Arnold SJ. Performance surfaces and adaptive landscapes. *Integrative and Comparative Biology* 2003;**43**:367–75. https://doi.org/10.1093/icb/43.3.367

Bapst DW. paelotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution* 2012;**3**:803–7. https://doi.org/10.1111/j.2041-210x.2012.00223.x

Barghusen HR. A review of fighting adaptations in dinocephalians (Reptilia, Therapsida). *Paleobiology* 1975;**1**:295–311. https://doi.org/10.1017/s0094837300002542

Benoit J, Manger PR, Norton L, *et al*. Synchrotron scanning reveals the palaeoneurology of the head-butting *Moschops capensis* (Therapsida, Dinocephalia). *PeerJ* 2017;**5**:e3496. https://doi.org/10.7717/peerj.3496

Berman DS, Sumida SS, Lombard RE. Reinterpretation of the temporal and occipital regions in *Diadectes* and the relationships of diadectomorphs. *Journal of Paleontology* 1992;**66**:481–99. https://doi.org/10.1017/s0022336000034028

Berman DS, Sumida SS, Martens T. *Diadectes* (Diadectomorpha: Diadectidae) from the Early Permian of central Germany, with description of a new species. *Annals of the Carnegie Museum* 1998;**67**:53–93. https://doi.org/10.5962/p.215205

Boitsova EA, Skutschas PP, Sennikov AG, *et al*. Bone histology of two pareiasaurs from Russia (*Deltavjatia rossica and Scutosaurus karpinskii*) with implications for pareiasaurian palaeobiology. *Biological Journal of the Linnean Society* 2019;**128**:289–310.

Bonett RM, Blair AL. Evidence for complex life cycle constraints on salamander body form diversification. *Proceedings of the National Academy of Sciences of the United States of America* 2017;**114**:9936–41. https://doi.org/10.1073/pnas.1703877114

Boos AD, Kammerer CF, Schultz CL, *et al*. A new dicynodont (Therapsida: Anomodontia) from the Permian of southern Brazil and its implications for bidentalian origins. *PLoS One* 2016;**11**:e0155000. https://doi.org/10.1371/journal.pone.0155000

Callier V, Clack JA, Ahlberg PE. Contrasting developmental trajectories in the earliest known tetrapod forelimbs. *Science* 2009;**324**:364–7. https://doi.org/10.1126/science.1167542

Canoville A, Chinsamy A. Bone microstructure of pareiasaurs (Parareptilia) from the Karoo Basin, South Africa: implications for growth strategies and lifestyle habits. *Anatomical Record (Hoboken, N. J.: 2007)* 2017;**300**:1039–66. https://doi.org/10.1002/ar.23534

Canoville A, Thomas DB, Chinsamy A. Insights into the habitat of Middle Permian pareiasaurs (Parareptilia) from preliminary isotopic analyses. *Lethaia* 2014;**47**:266–74. https://doi.org/10.1111/let.12056

Case EC. *A Revision of the Cotylosauria of North America*. Washington, DC: Carnegie Institute of Washington, 1911.

Clack JA. *Gaining Ground: The Origin and Evolution of Tetrapods*. Bloomington, IN: Indiana University Press, 2012.

Coatham SJ, Vinther J, Rayfield EJ, *et al*. Was the Devonian placoderm *Titanichthys* a suspension feeder? *Royal Society Open Science* 2020;**7**:200272. https://doi.org/10.1098/rsos.200272

Cohen KM, Finney SC, Gibbard PL, *et al*. The ICS International chronostratigraphic chart. *Episodes* 2013;**36**:199–204.

Cucherousset J, Boulêtreau S, Azémar F, *et al*. 'Freshwater killer whales': beaching behavior of an alien fish to hunt land birds. *PLoS One* 2012;**7**:e50840. https://doi.org/10.1371/journal.pone.0050840

Daeschler EB, Shubin NH, Jenkins FA Jr. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature* 2006;**440**:757–63. https://doi.org/10.1038/nature04639

Deakin WJ, Anderson PSL, den Boer W, *et al*. Increasing morphological disparity and decreasing optimality for jaw speed and strength during the radiation of jawed vertebrates. *Science Advances* 2022;**8**:eabl3644. https://doi.org/10.1126/sciadv.abl3644

Dickson BV, Pierce SE. Functional performance of turtle humerus shape across an ecological adaptive landscape. *Evolution* 2019;**73**:1265–77. https://doi.org/10.1111/evo.13747

Dickson BV, Clack JA, Smithson TR, *et al*. Functional adaptive landscapes predict terrestrial capacity at the origin of limbs. *Nature* 2021;**589**:242–5. https://doi.org/10.1038/s41586-020-2974-5

Didier G, Laurin M. Distributions of extinction times from fossil ages and tree topologies: the example of mid-Permian synapsid extinctions. *PeerJ* 2021;**9**:e12577. https://doi.org/10.7717/peerj.12577

Didier G, Laurin M. Testing extinction events and temporal shifts in diversification and fossilization rates through the skyline Fossilized Birth-Death (FBD) model: The example of some mid-Permian synapsid extinctions. *Cladistics* 2024;**40**:282–306. https://doi.org/10.1111/cla.12577

Eberth DA, Berman DS, Sumida SS, *et al*. Lower Permian terrestrial paleoenvironments and vertebrate paleoecology of the Tambach Basin (Thuringia, central Germany): the upland holy grail. *Palaios* 2000;**15**:293–313. https://doi.org/10.2307/3515538

Esteve-Altava B, Marugán-Lobón J, Botella H, *et al*. Structural constraints in the evolution of the tetrapod skull complexity: Williston's law

revisited using network models. *Evolutionary Biology* 2013;**40**:209–19. https://doi.org/10.1007/s11692-012-9200-9
Fabre A-C, Bardua C, Bon M, *et al*. Metamorphosis shapes cranial diversity and rate of evolution in salamanders. *Nature Ecology & Evolution* 2020;**4**:1129–40. https://doi.org/10.1038/s41559-020-1225-3
Ferrón HG, Martínez-Pérez C, Rahman IA, *et al*. Computational fluid dynamics suggests ecological diversification among stem-gnathostomes. *Current Biology* 2020;**30**:4808–4813. e4803.
Fletcher TM, Janis CM, Rayfield EJ. Finite element analysis of ungulate jaws: can mode of digestive physiology be determined? *Palaeontologia Electronica* 2010;**13**:21A:15p.
Fonseca CM, Fleming PJ. Multiobjective optimization and multiple constraint handling with evolutionary algorithms. I. A unified formulation. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans* 1998;**28**:26–37. https://doi.org/10.1109/3468.650319
Ford DP, Benson RB. The phylogeny of early amniotes and the affinities of Parareptilia and Varanopidae. *Nature Ecology & Evolution* 2020;**4**:57–65.
Fortuny J, Marcé-Nogué J, De Esteban-Trivigno S, *et al*. Temnospondyli bite club: ecomorphological patterns of the most diverse group of early tetrapods. *Journal of Evolutionary Biology* 2011;**24**:2040–54. https://doi.org/10.1111/j.1420-9101.2011.02338.x
Frazzetta T. Adaptive problems and possibilities in the temporal fenestration of tetrapod skulls. *Journal of Morphology* 1968;**125**:145–57.
Gill PG, Purnell MA, Crumpton N, *et al*. Dietary specializations and diversity in feeding ecology of the earliest stem mammals. *Nature* 2014;**512**:303–5. https://doi.org/10.1038/nature13622
Goldberg DE. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc., 1989.
Guillerme T, Cooper N. Time for a rethink: time sub-sampling methods in disparity-through-time analyses. *Palaeontology* 2018;**61**:481–93. https://doi.org/10.1111/pala.12364
Guillerme T, Cooper N, Brusatte SL, *et al*. Disparities in the analysis of morphological disparity. *Biology Letters* 2020a;**16**:20200199.
Guillerme T, Puttick MN, Marcy AE, *et al*. Shifting spaces: Which disparity or dissimilarity measurement best summarize occupancy in multidimensional spaces? *Ecology and Evolution* 2020b;**10**:7261–75.
Gutarra S, Rahman IA. The locomotion of extinct secondarily aquatic tetrapods. *Biological Reviews of the Cambridge Philosophical Society* 2022;**97**:67–98. https://doi.org/10.1111/brv.12790
Heiss E, Aerts P, Van Wassenbergh S. Masters of change: seasonal plasticity in the prey-capture behavior of the Alpine newt *Ichthyosaura alpestris* (Salamandridae). *The Journal of Experimental Biology* 2013;**216**:4426–34. https://doi.org/10.1242/jeb.091991
Heiss E, Aerts P, Van Wassenbergh S. Aquatic–terrestrial transitions of feeding systems in vertebrates: a mechanical perspective. *Journal of Experimental Biology* 2018;**221**:20180425.
Hill JJ, Puttick MN, Stubbs TL, *et al*. Evolution of jaw disparity in fishes. *Palaeontology* 2018;**61**:847–54. https://doi.org/10.1111/pala.12371
Horn J, Nafpliotis N, Goldberg DE. A niched Pareto genetic algorithm for multiobjective optimization. In: *Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence: IEEE*, 1994, 82–7.
Hotton N, Olson EC, Beerbower R, *et al. Amniote Origins and the Discovery of Herbivory. Amniote Origins*. San Diego, CA: Academic Press, 1997, 207–64.
Janis CM, Keller JC. Modes of ventilation in early tetrapods: costal aspiration as a key feature of amniotes. *Acta Palaeontologica Polonica* 2001;**46**:20010101.
Jones KE, Dickson BV, Angielczyk KD, *et al*. Adaptive landscapes challenge the 'lateral-to-sagittal' paradigm for mammalian vertebral evolution. *Current Biology* 2021;**31**:1883–92.e7. https://doi.org/10.1016/j.cub.2021.02.009
Karlsruhe WM, Sues H-D. Gut contents of *Parasaurus* (Pareiasauria) and *Protorosaurus* (Archosauromorpha) from the Kupferschiefer (Upper Permian) of Hessen, Germany. *Paläontologische Zeitschrift* 1993;**67**:169–76. https://doi.org/10.1007/bf02985876
Klembara J, Hain M, Čerňanský A, *et al*. Anatomy of the neural endocranium, parasphenoid and stapes of *Diadectes absitus* (Diadectomorpha) from the early Permian of Germany based on the high-resolution X-ray microcomputed tomography. *The Anatomical Record* 2020a;**303**:2977–99. https://doi.org/10.1002/ar.24376
Klembara J, Hain M, Ruta M, *et al*. Inner ear morphology of diadectomorphs and seymouriamorphs (Tetrapoda) uncovered by high-resolution x-ray microcomputed tomography, and the origin of the amniote crown group. *Palaeontology* 2020b;**63**:131–54. https://doi.org/10.1111/pala.12448
Kligman BT, Gee BM, Marsh AD, *et al*. Triassic stem caecilian supports dissorophoid origin of living amphibians. *Nature* 2023;**614**:102–7. https://doi.org/10.1038/s41586-022-05646-5
Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing* 1982;**18**:236–58. https://doi.org/10.1016/0146-664x(82)90034-x
Lee MSY. Pareiasaur phylogeny and the origin of turtles. *Zoological Journal of the Linnean Society* 1997;**120**:197–280. https://doi.org/10.1111/j.1096-3642.1997.tb01279.x
Lemberg JB, Daeschler EB, Shubin NH. The feeding system of *Tiktaalik roseae*: an intermediate between suction feeding and biting. *Proceedings of the National Academy of Sciences of the United States of America* 2021;**118**:e2016421118. https://doi.org/10.1073/pnas.2016421118
Long JA, Gordon MS. The greatest step in vertebrate history: a paleobiological review of the fish–tetrapod transition. *Physiological and Biochemical Zoology: PBZ* 2004;**77**:700–19. https://doi.org/10.1086/425183
Ma W, Pittman M, Butler RJ, *et al*. Macroevolutionary trends in theropod dinosaur feeding mechanics. *Current Biology: CB* 2022;**32**:677–86.e3. https://doi.org/10.1016/j.cub.2021.11.060
Marcé-Nogué J, Püschel TA, Daasch A, *et al*. Broad-scale morphofunctional traits of the mandible suggest no hard food adaptation in the hominin lineage. *Scientific Reports* 2020;**10**:6793. https://doi.org/10.1038/s41598-020-63739-5
Marjanović D, Laurin M. Phylogeny of Paleozoic limbed vertebrates reassessed through revision and expansion of the largest published relevant data matrix. *PeerJ* 2019;**6**:e5565. https://doi.org/10.7717/peerj.5565
Marjanović D, Maddin HC, Olori JC, *et al*. The new problem of *Chinlestegophis* and the origin of caecilians (Amphibia, Gymnophionomorpha) is highly sensitive to old problems of sampling and character construction. *Fossil Record* 2024;**27**:55.
Markey MJ, Marshall CR. Terrestrial-style feeding in a very early aquatic tetrapod is supported by evidence from experimental analysis of suture morphology. *Proceedings of the National Academy of Sciences of the United States of America* 2007;**104**:7134–8. https://doi.org/10.1073/pnas.0701706104
McGhee GR. Theoretical morphology: the concept and its applications. *Short Courses in Paleontology* 1991;**4**:87–102. https://doi.org/10.1017/s2475263000002130
McGhee GR. *The Geometry of Evolution: Adaptive Landscapes and Theoretical Morphospaces*. Cambridge: Cambridge University Press, 2006
McGhee GR. Limits in the evolution of biological form: a theoretical morphologic perspective. *Interface Focus* 2015;**5**:20150034. https://doi.org/10.1098/rsfs.2015.0034
Michel KB, Adriaens D, Aerts P, *et al*. Functional anatomy and kinematics of the oral jaw system during terrestrial feeding in *Periophthalmus barbarus*. *Journal of Morphology* 2014;**275**:1145–60. https://doi.org/10.1002/jmor.20291
Michel KB, Aerts P, Gibb AC, *et al*. Functional morphology and kinematics of terrestrial feeding in the largescale foureyes (*Anableps anableps*). *The Journal of Experimental Biology* 2015a;**218**:2951–60. https://doi.org/10.1242/jeb.124644
Michel KB, Heiss E, Aerts P, *et al*. A fish that uses its hydrodynamic tongue to feed on land. *Proceedings Biological Sciences* 2015b;**282**:20150057. https://doi.org/10.1098/rspb.2015.0057

Morales-García NM, D BT, Hill JJ, *et al.* The use of extruded finite-element models as a novel alternative to tomography-based models: a case study using early mammal jaws. *Journal of Royal Society Interface* 2019;**16**:20190674.

Morris ZS, Abzhanov A. Heading for higher ground: developmental origins and evolutionary diversification of the amniote face. *Current Topics in Developmental Biology* 2021;**141**:241–77. https://doi.org/10.1016/bs.ctdb.2020.12.003

Nakaya K. Hydrodynamic function of the head in the hammerhead sharks (Elasmobranchii: Sphyrnidae). *Copeia* 1995;**1995**:330–6. https://doi.org/10.2307/1446895

Neenan JM, Ruta M, Clack JA, *et al.* Feeding biomechanics in *Acanthostega* and across the fish–tetrapod transition. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 2014;**281**:20132689.

Olson EC. Jaw mechanisms: rhipidistians, amphibians, reptiles. *American Zoologist* 1961;**1**:205–15. https://doi.org/10.1093/icb/1.2.205

Palci A, Lee MS, Hutchinson MN. Patterns of postnatal ontogeny of the skull and lower jaw of snakes as revealed by micro-CT scan data and three-dimensional geometric morphometrics. *Journal of Anatomy* 2016;**229**:723–54. https://doi.org/10.1111/joa.12509

Panagiotopoulou O. Finite element analysis (FEA): applying an engineering method to functional morphology in anthropology and human biology. *Annals of Human Biology* 2009;**36**:609–23. https://doi.org/10.1080/03014460903019879

Pardo JD, Szostakiwskyj M, Ahlberg PE, *et al.* Hidden morphological diversity among early tetrapods. *Nature* 2017;**546**:642–5. https://doi.org/10.1038/nature22966

Polly PD, Stayton CT, Dumont ER, *et al.* Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis. *Journal of Vertebrate Paleontology* 2016;**36**:e1111225. https://doi.org/10.1080/02724634.2016.1111225

Porro LB, Rayfield EJ, Clack JA. Computed tomography, anatomical description and three-dimensional reconstruction of the lower jaw of *Eusthenopteron foordi* Whiteaves, 1881 from the Upper Devonian of Canada. *Palaeontology* 2015a;**58**:1031–47. https://doi.org/10.1111/pala.12192

Porro LB, Rayfield EJ, Clack JA. Descriptive anatomy and three-dimensional reconstruction of the skull of the early tetrapod *Acanthostega gunnari* Jarvik, 1952. *PLoS One* 2015b;**10**:e0118882. https://doi.org/10.1371/journal.pone.0118882

Powder KE, Milch K, Asselin G, *et al.* Constraint and diversification of developmental trajectories in cichlid facial morphologies. *EvoDevo* 2015;**6**:1–14.

Raup DM. Geometric analysis of shell coiling in ammonoids. *Journal of Paleontology* 1967;**41**:43–65.

Raup DM, Michelson A. Theoretical morphology of the coiled shell. *Science* 1965;**147**:1294–5. https://doi.org/10.1126/science.147.3663.1294

Rawson JR, Esteve-Altava B, Porro LB, *et al.* Early tetrapod cranial evolution is characterized by increased complexity, constraint, and an offset from fin-limb evolution. *Science Advances* 2022;**8**:eadc8875.

Rayfield EJ. Finite element analysis and understanding the biomechanics and evolution of living and fossil organisms. *Annual Review of Earth and Planetary Sciences* 2007;**35**:541–76. https://doi.org/10.1146/annurev.earth.35.031306.140104

Reisz RR, Maho T, Modesto SP. Recumbirostran 'microsaurs' are not amniotes. *Journal of Systematic Palaeontology* 2024;**22**:2296078.

Romano M, Nicosia U. Cladistic analysis of Caseidae (Caseasauria, Synapsida): using the gap-weighting method to include taxa based on incomplete specimens. *Palaeontology* 2015;**58**:1109–30. https://doi.org/10.1111/pala.12197

San Mauro D, Agorreta A, Garcia-Porta J. Origin, evolution and diversification of extant amphibians. In: *Evolutionary Ecology of Amphibians*. Boca Raton, FL: CRC Press, 2023, 7–23.

Schaeffer J, Benton MJ, Rayfield EJ, *et al.* Morphological disparity in theropod jaws: comparing discrete characters and geometric morphometrics. *Palaeontology* 2019;**63**:283–99. https://doi.org/10.1111/pala.12455

Schoch RR. Evolution of life cycles in early amphibians. *Annual Review of Earth and Planetary Sciences* 2009;**37**:135–62. https://doi.org/10.1146/annurev.earth.031208.100113

Schoch RR. The evolution of major temnospondyl clades: an inclusive phylogenetic analysis. *Journal of Systematic Palaeontology* 2013;**11**:673–705. https://doi.org/10.1080/14772019.2012.699006

Seilacher A. Arbeitskonzept zur konstruktions-morphologie. *Lethaia* 1970;**3**:393–6. https://doi.org/10.1111/j.1502-3931.1970.tb00830.x

Snively E, Anderson PSL, Ryan MJ. Functional and ontogenetic implications of bite stress in arthrodire placoderms. *Kirtlandia* 2010;**57**:53–60.

Stayton CT. Terrestrial feeding in aquatic turtles: environment-dependent feeding behavior modulation and the evolution of terrestrial feeding in Emydidae. *The Journal of Experimental Biology* 2011;**214**:4083–91. https://doi.org/10.1242/jeb.060574

Stayton CT. Performance in three shell functions predicts the phenotypic distribution of hard-shelled turtles. *Evolution* 2019a;**73**:720–34. https://doi.org/10.1111/evo.13709

Stayton CT. Performance surface analysis identifies consistent functional patterns across 10 morphologically divergent terrestrial turtle lineages. *Integrative and Comparative Biology* 2019b;**59**:346–57. https://doi.org/10.1093/icb/icz072

Sues H-D, Reisz RR. Origins and early evolution of herbivory in tetrapods. *Trends in Ecology & Evolution* 1998;**13**:141–5. https://doi.org/10.1016/s0169-5347(97)01257-3

Svensson E, Calsbeek R. *The Adaptive Landscape in Evolutionary Biology*. Oxford: Oxford University Press, 2012.

Swartz B. A marine stem-tetrapod from the Devonian of western North America. *PLoS One* 2012;**7**:e33683. https://doi.org/10.1371/journal.pone.0033683

Tsuji LA. Anatomy, cranial ontogeny and phylogenetic relationships of the pareiasaur *Deltavjatia rossicus* from the Late Permian of central Russia. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 2013;**104**:81–122. https://doi.org/10.1017/s1755691013000492

Tsuji LA, Müller J. Assembling the history of the Parareptilia: phylogeny, diversification, and a new definition of the clade. *Fossil Record* 2009;**12**:71–81. https://doi.org/10.5194/fr-12-71-2009

Van den Brandt MJ, Day MO, Manucci F, *et al.* First volumetric body mass estimate and a new in vivo 3D reconstruction of the oldest Karoo pareiasaur *Bradysaurus baini*, and body size evolution in Pareiasauria. *Historical Biology* 2023;**36**:587–601. https://doi.org/10.1080/08912963.2023.2175211

Van Wassenbergh S. *Kinematics of Terrestrial Capture of Prey by the Eelcatfish Channallabes apus*. Oxford: Oxford University Press, 2013.

Van Wassenbergh S, Herrel A, Adriaens D, *et al.* A catfish that can strike its prey on land. *Nature* 2006;**440**:881–881. https://doi.org/10.1038/440881a

Van Wassenbergh S, Brecko J, Aerts P, *et al.* Hydrodynamic constraints on prey-capture performance in forward-striking snakes. *Journal of the Royal Society Interface* 2009;**7**:773–85. https://doi.org/10.1098/rsif.2009.0385

Van Wassenbergh S, Potes NZ, Adriaens D. Hydrodynamic drag constrains head enlargement for mouthbrooding in cichlids. *Journal of the Royal Society Interface* 2015;**12**:20150461. https://doi.org/10.1098/rsif.2015.0461

Van Wassenbergh S, Bonte C, Michel KB. Terrestrial capture of prey by the reedfish, a model species for stem tetrapods. *Ecology and Evolution* 2017;**7**:3856–60. https://doi.org/10.1002/ece3.2694

Villamil J, Demarco PN, Meneghel M, *et al.* Optimal swimming speed estimates in the Early Permian mesosaurid *Mesosaurus tenuidens* (Gervais 1865) from Uruguay. *Historical Biology* 2016;**28**:963–71.

Witzmann F. Phylogenetic patterns of character evolution in the hyobranchial apparatus of early tetrapods. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 2013;**104**:145–67. https://doi.org/10.1017/s1755691013000480

Wright S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the XI International Congress of Genetics* 1932;**8**:209–22.